

WHAT IS CLAIMED IS:

Sub
A3

00000000000000000000

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

1. A method for operating a network of point of presence servers .
sharing a hostname, the method comprises:
receiving a request from a user for a web page at a first web address, the first
web address including the hostname;
determining traffic loads of a plurality of customer web servers, each of the
customer web servers storing the web page;
determining a customer web server from the plurality of customer web servers
that is appropriate for the request, the customer web server having a traffic load lower than
traffic loads of remaining customer web servers from the plurality of customer web servers;
determining an IP address of the customer web server;
directing the request from the user to the customer web server; thereafter
receiving a request from the user for static content on the web page at a
second web address, the second web address including the hostname;
determining service metrics of point of presence servers in the network of
point of presence servers;
determining the point of presence server from the network of point of
presence servers that is appropriate for the request, the point of presence server having service
metrics better than service metrics of remaining point of presence servers from the network of
point of presence servers;
retrieving the static content from the point of presence server; and
providing the static content to the user.

1
2
3
4
5
6

2. The method of claim 1 further comprising determining load of
point of presence servers in the network of point of presence servers;
wherein determining the point of presence server from the network of point of
presence servers that is appropriate for the request, the point of presence server having a
latency and a load lower than latency or load of the remaining point of presence servers from
the network of point of presence servers.

1
2
3
4

3. The method of claim 1 further comprising:
determining whether the point of presence server includes the static content;
determining a web server that includes the static content when the point of
presence server does not include the static content;

5 retrieving the static content from the web server that includes the static
 6 content; and
 7 storing the static content from the web server in the point of presence server.

1 4. The method of claim 3 wherein determining the web server
 2 comprises:
 3 determining traffic loads of the plurality of customer web servers, each of the
 4 customer web servers storing the static content; and
 5 determining another customer web server from the plurality of customer web
 6 servers that is appropriate for the request, the another customer web server having a traffic
 7 load lower than traffic loads of remaining customer web servers from the plurality of
 8 customer web servers.

1 5. The method of claim 4 wherein retrieving the static content from
 2 the web server comprises:
 3 determining another IP address of the another customer web server; and
 4 requesting the static content from the another customer web server at the
 5 another IP address.

1 6. The method of claim 1 wherein the network of point of presence
 2 servers comprises a domain name server.

1 7. The method of claim 1
 2 wherein the request from the user for the web page is transferred from a first
 3 domain name server;
 4 wherein the network of point of presence servers comprises a second domain
 5 name server; and
 6 wherein the second domain name server determines the customer web server
 7 from the plurality of customer web servers.

1 8. A method for operating a network of point of presence servers
 2 comprises:
 3 receiving a first request from a client DNS server to resolve a first domain
 4 name, the client DNS server receiving a request from a user of a web page address that
 5 includes the first domain name;

6 determining load measurements of a plurality of customer web servers, each
7 of the customer web servers addressable by the first domain name, and each of the customer
8 web servers configured to service the request from the user;
9 determining a customer web server from the plurality of customer web
10 servers, the customer web server having load measurements lower than load measurements of
11 other customer web servers from the plurality of customer web servers;
12 determining an IP address of the customer web server;
13 providing the IP address of the customer web server to the client DNS server;
14 thereafter
15 receiving a second request from the client DNS server to resolve a second
16 domain name, the client DNS server receiving a request from the user of a uniform resource
17 locator that includes the second domain name;
18 determining performance metric measurement of point of presence servers in
19 the network of point of presence servers, each of the point of presence servers addressable by
20 the second domain name;
21 determining a point of presence server from the network of point of presence
22 servers, the point of presence server having performance metric measurements lower than
23 performance metric measurements of other point of presence servers from the network of
24 point of presence servers;
25 providing the IP address of the point of presence server to the client DNS
26 server;
27 retrieving data from the point of presence server in response to the uniform
28 resource locator; and
29 providing the data to the user.

1 9. The method of claim 8 wherein the load measurements comprise
2 latency measurements.

1 10. The method of claim 8 wherein the performance metric
2 measurements comprise load CPU and memory measurements, HTTP response
3 measurements, FTP response measurements.

1 11. The method of claim 8 wherein retrieving data from the point of
2 presence server comprises:
3 determining whether the point of presence server includes the data;

1 15. A method for a network of point of presence servers comprises:
2 receiving at a first point of presence server a first request from a first client
3 DNS server to resolve a first domain name, the first request from the first client DNS server
4 determined in response to a first uniform resource locator entered by a first user, the first
5 uniform resource locator comprising the first domain name;
6 receiving at a second point of presence server a first request from a second
7 client DNS server to resolve the first domain name, the first request from the second client
8 DNS server determined in response to the first uniform resource locator entered by a second
9 user, the first uniform resource locator comprising the first domain name;
10 determining at the first point of presence server traffic measurements of a
11 plurality of customer web servers, each of the customer web servers addressable by the first
12 domain name, and each of the customer web servers storing data associated with the first
13 uniform resource locator;
14 determining at the first point of presence server a first customer web server
15 from the plurality of customer web servers, the first customer web server having traffic
16 measurements lower than load measurements of other customer web servers from the
17 plurality of customer web servers;
18 determining at the second point of presence server the first customer web
19 server from the plurality of customer web servers, the first customer web server having
20 traffic measurements lower than load measurements of other customer web servers from the
21 plurality of customer web servers;
22 determining at the first point of presence server an IP address of the first
23 customer web server;
24 determining at the second point of presence server an IP address of the first
25 customer web server;
26 providing from the first point of presence server the IP address of the first
27 customer web server to the first client DNS server and to the second client DNS server;
28 providing from the second point of presence server the IP address of the first
29 customer web server to the first client DNS server and to the second client DNS server;
30 thereafter
31 receiving at the first point of presence server a second request from the first
32 client DNS server to resolve a second domain name, the second request from the first client
33 DNS server determined in response to a second uniform resource locator comprising the
34 second domain name, the second uniform resource locator from the first customer web server;

